09876235

attach #17

VERIFICATION OF TRANSLATION

IN THE NAME OF :

Japanese Patent Application No. 8-274855
Filed on October 17, 1996

I, Yoshiyuki KAWAGUCHI, a citizen of Japan and the translator of the document attached, whose address is C/O SERA, TOYAMA, MATSUKURA & KAWAGUCHI, Yokoyama Building 6th Floor, 4-10, Higashi Nihonbashi 3-chome, Chuo-Ku, Tokyo, Japan, state that the following is a true translation of the Japanese Patent Application No. 8-274855 filed on October 17, 1996 to the best of my knowledge and belief.

Signed at Tokyo, Japan.
This is 15th day of June, 2001

Y Kawaguchi

Yoshiyuki KAWAGUCHI
Patent Attorney
SERA, TOYAMA, MATSUKURA & KAWAGUCHI

APPENDIX 2

[Name of Document] Application for Patent

[Reference Number] 961542

[Date of Filing] October 17, 1996

[Addressee] Commissioner of the Patent Office

[Int. Class.] C12N 15/00

[Title of the Invention] Molecule Assigning Genotype to Phenotype

[Number of Claims] 11

[Inventor]

[Address] c/o Mitsubishi Kasei Institute of Life Sciences, 11, Minamiooya, Machida-shi, Tokyo-to

[Name] Hiroshi Yanagawa

[Inventor]

[Address] c/o Mitsubishi Kasei Institute of Life Sciences, 11, Minamiooya, Machida-shi, Tokyo-to

[Name] Naoto Nemoto

[Inventor]

[Address] 1-18-18, Teraya, Tsurumi-ku, Yokohama-shi, Kanagawa-ken

[Name] Etsuko Miyamoto

[Inventor]

[Address] 671-6, Zinde, Urawa-shi, Saitama-ken

[Name] Yuzuru Fusimi

[Applicant]

[Identification Number] 000005968

[Name] Mitsubishi Chemical Corporation

[Representative] Akira Miura

[Agent]

[Identification Number] 100103997

[Patent Attorney]

[Name] Koji Hasegawa

[Fee]

[Ledger Number] 035035

[Amount] 21000

[List of Attached Items]

[Item Name] Specification 1

[Item Name] Drawings 1

[Item Name] Abstract 1

[Number of General Power of Attorney] 9306391

[Name of Document] Specification

[Title of the Invention] Molecule Assigning Genotype to Phenotype

[Claims]

1. A molecule assigning a genotype to a phenotype, which comprises a nucleic acid portion composed of a polymer of nucleic acid, and a protein portion.

2. The assigning molecule according to claim 1, wherein a 3'-terminal end of the nucleic acid portion and a C-terminal end of the protein portion are bonded with a covalent bond.

3. The assigning molecule according to claim 1 or 2, wherein a 3'-terminal end of the nucleic acid portion covalently bonded to a C-terminal end of the protein portion is puromycin.

4. The assigning molecule according to any one of claims 1 to 3, wherein the nucleic acid portion comprises a gene, and the protein portion is a translation product of the gene of the nucleic acid portion.

5. The assigning molecule according to any one of claims 1 to 3, wherein the nucleic acid portion comprises a gene composed of RNA, a spacer and a suppressor tRNA, and the protein portion is a translation product of the gene of the nucleic acid portion.

6. The assigning molecule according to any one of claims 1 to 3, wherein the RNA of the nucleic acid portion comprises an initiation codon and a termination codon, wherein the suppressor tRNA comprises an anticodon corresponding to the

termination codon (umber), and wherein the protein portion is a translation product of the RNA gene of the nucleic acid portion.

7. The assigning molecule according to any one of claims 1 to 3, wherein the nucleic acid portion comprises a gene composed of RNA, and a spacer composed of DNA and RNA, and wherein the protein portion is a translation product of the gene of the nucleic acid portion.

8. The assigning molecule according to any one of claims 1 to 3, wherein the nucleic acid portion comprises a gene composed of DNA, and a spacer composed of DNA and RNA, and wherein the protein portion is a translation product of the gene of the nucleic acid portion.

9. A method for constructing an assigning molecule, which comprises (a) bonding a DNA comprising a sequence corresponding to a suppressor tRNA, to a 3'-terminal end of a DNA (gene) through a spacer, (b) transcribing the obtained DNA bonded product into RNA, (c) bonding, to a 3'-terminal end thereof, a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, which can be covalently bound to an amino acid or a substance having a chemical structure analogous to that of an amino acid, and (d) performing protein synthesis in a cell-free protein synthesis system using the obtained RNA-bonded product as a gene to connect a genotype to a phenotype.

10. A method for constructing an assigning molecule, which comprises (a) preparing a DNA (gene) which has no termination

codon, (b) transcribing the prepared DNA into RNA (gene), (c) bonding a chimeric spacer composed of DNA and RNA to a 3'-terminal end thereof, (d) bonding, to a 3'-terminal end thereof, a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, which can be covalently bound to an amino acid or a substance having a chemical structure analogous to that of an amino acid, and (e) performing protein synthesis in an *E. coli* cell-free protein synthesis system using the obtained RNA-DNA chimeric bonded product as a gene to connect a genotype to a phenotype.

11. The construction method according to claim 9 or 10, wherein the nucleic acid or the substance having the chemical structure analogous to that of the nucleic acid and a base is puromycin.

[Description of the Invention]

[0001]

[Technical Field to Which the Invention Pertains]

The present invention relates to a molecule assigning a genotype to a phenotype. More specifically, it relates to a molecule assigning a genotype to a phenotype, comprising a nucleic acid portion composed of a polymer of nucleic acid and a protein portion composed of a polymer of amino acid. The assigning molecule of the present invention is a highly useful substance that can be utilized in evolutionary molecular engineering such as in the modification of functional biopolymers such enzymes, antibodies and ribozymes and

creation of biopolymers having functions not found in living organisms.

[0002]

Through advances in biochemistry, molecular biology and biophysics, it has been learned that living organisms are molecular machines which function and propagate by interactions among molecules. Among the characteristics of earth's living organisms, the fundamentals are their preservation of genetic information in DNA nucleotide sequences and their ability to translate this information into functional proteins through the medium of mRNA. Owing to progress in genetic engineering, biopolymers with given sequences, like nucleotides and peptides, can now be easily synthesized. Protein engineering and RNA engineering, today a focus of attention, owe their existence to genetic engineering. The aims of protein engineering and RNA engineering are to solve the puzzle of the three-dimensional structures required for proteins and RNA fulfilling specific functions and to enable humans to freely design proteins and RNA possessing desired functions. Because of the diversity and complexity of these structures and the difficulty of a theoretical approach to their three-dimensional structures, however, current protein engineering and RNA engineering are still at the stage of modifying some of residues at active sites and observing changes in the structure and functions. Human knowledge has thus not yet reached the stage of designing proteins and RNA.

[0003]

Understanding the functions of biopolymers in their relationship to the elemental processes of higher life phenomena will require elucidation of the correlation between protein molecular structure and function. The line of thought we take in the following is not only to make the best of "human knowledge" but also to take advantage of the "wisdom of nature." This is because we concluded that we would have to acquire the ability to put both to work in order to overcome the current difficulties of protein engineering and move forward with the design and production of functional biopolymers. When the classical methods are diverted to the design of proteins with new functions and activities, the difficulty of protein design by site-specific mutations can sometimes be avoided. This can be called "taking advantage of the wisdom of nature."

[0004]

Although the drawback of this method is the difficulty of screening to identify mutants with new functions and activities, this difficulty is overcome by the RNA catalysts that have recently come into the spotlight. Attempts have been made to select an RNA with specific characteristics from among RNAs synthesized to have an extremely large number of random sequences (about 10^{13} types) (Ellington, A. D. & Szostak, J. W. (1990) Nature, 346, 818-822).

[0005]

This is an example of evolutionary molecular engineering. As typified by this example, the primary goal in the

evolutionary molecular engineering of proteins is to find out optimum sequences by searching an expansive sequence space of a scale unimaginable in conventional protein engineering. By "making the best of human knowledge" to devise a screening system for this, it will be possible to discover numerous quasi-optimum sequences around the optimum sequences and thus to construct an experimental system for studying "sequence vs function."

[0006]

The remarkable functions of living bodies were acquired through the process of evolution. Therefore, if evolution can be replicated, it should be possible to modify enzymes, antibodies, ribozymes and other functional biopolymers and, further, to create biopolymers with functions not found in living organisms in the laboratory. Needless to say, research on protein modification and creation is an object of utmost importance to various aspects of biotechnology such as utilization of enzymes as industrial catalysts, biochips, biosensors and sugar-chain engineering.

[0007]

Given the fact that molecular design utilizing structural theory is, as symbolized by the continuing high regard for "screening," still in an unperfected state, the evolutionary technique has a practical value for utilization in selecting useful proteins as a more efficient strategy. Building a "time machine" capable of more efficiently producing evolution in a laboratory, if such were possible, would not

only enable modification of enzymes, antibodies (vaccines, monoclonal antibodies etc.) and other existing proteins but also open the way to the creation of enzymes for decomposing environmental contaminants, purifiers and others and new proteins not present in the biological world. If an experimental system for protein evolution can be established, therefore, it can be expected to be aggressively utilizable for application in a wide range of fields including power saving and energy preservation in industrial processes, energy production and environmental preservation. The assigning molecule of the present invention is a highly useful substance in protein modification and other aspects of evolutionary molecular engineering.

[0008]

[Prior Art]

Evolutionary molecular engineering is a field of study that attempts to conduct molecular design of functional polymers by utilizing high-speed molecular evolution in the laboratory, i.e., by laboratory investigation and optimization of the adaptive locomotion of biopolymers in sequence space. It is a completely new molecular biotechnology that first produced substantial results in 1990 (Yuzuru Husimi (1991) Kagaku, 61, 333-340; Yuzuru Husimi (1992) Koza Shinka, Vol. 6, University of Tokyo Publishing Society).

[0009]

Life is a product of molecular evolution and natural selection. The evolution of molecules is a universal life

phenomenon but its mechanism is not something that can be elucidated by studies that track the history of past evolution. Rather, the approach of constructing and studying the behavior of simple molecules and life systems that evolve in the laboratory better provides fundamental knowledge regarding molecular evolution and enables establishment of a verifiable theory applicable in molecular engineering.

[0010]

It is known that a polymer system will evolve if it satisfies the following five conditions: (1) an open system far out of equilibrium, (2) a self-replicative system, (3) a mutation system, (4) a system with genotype and phenotype assignment strategy, and (5) a system with appropriate adaptation topography in sequence space. (1) and (2) are conditions for occurrence of natural selection and (5) is determined beforehand by the physicochemical properties of the biopolymer. The genotype and phenotype assignment of (4) is a prerequisite for evolution by natural selection.

[0011]

The following three strategies are adopted in both the natural world and evolutionary molecular engineering: (a) nucleic acid-type in which the genotype and the phenotype are carried on the same molecule, (b) virus-type in which the genotype and the phenotype are made to be one bonded product, and (c) a cell-type in which the genotype and the phenotype are contained in a single compartment (Figure 1).

As the nucleic acid-type (a) in which the genotype and

the phenotype are carried on the same molecule is a simple system, success with RNA catalysts (ribozymes) has already been reported (Hiroshi Yanagawa (1993) New Age of RNA, pp.55-77, Yodosha).

[0012]

Conceivable problem points of the cell-type (c) are (1) the averaging effect, (2) the eccentricity effect and (3) the random replication effect. The averaging effect arises because the assignment of the genotype to the phenotype statistically averages out and becomes ambiguous when the number of copies of the cell genome is large. Since an evolved genome is only one among the number of copies in a cell (n), performance enhancement averages out and a struggle for existence in the cell population begins at selection coefficient (s)/n. A smaller copy number (n) is therefore advantageous for the cell-type. Due to the presence of the eccentricity effect, however, when the number of segments is large, n must be very large to prevent the eccentricity effect. The apparent selection coefficient in the struggle for existence in the cell population can therefore be expected to be very much smaller than in the case of the virus-type. Since the time required for selection is proportional to the reciprocal of the selection coefficient, the rate of evolution is much slower than that of the virus-type. Further, the random replication effect (3) is fatal to the cell-type. This is because the random replication of segmented essential genes by this effect makes replication of all essential genes prior

to cell division extremely difficult. This means that even if an essential gene with an advantageous mutation should occur, the probability of its being replicated and passed on to a daughter cell is extremely low.

[0013]

Chemically bonding the genotype and the phenotype as in the virus-type (b) to unite them is necessary for efficient evolution.

Various techniques have already been proposed and are in the process of development for evolutionary molecular engineering of the virus-type (b), including phage display (Smith, G. P. (1985) Science 228, 1315-1317; Scott, J. K. & Smith, G. P. (1990) Science 249, 386-390), polysome display (Mattheakis, L. C. et al. (1994) Proc. Natl. Acad. Sci. USA 91, 9022-9026), encoded combinatorial library (Brenner, S. & Lerner, R. A. (1992) Proc. Natl. Acad. Sci. USA 89, 5381-5383), and cellstat (Husimi, Y. et al. (1982) Rev. Sci. Instrum. 53, 517-522).

Despite the importance of the magnitude of the searchable sequence space in evolutionary molecular engineering, however, a method for globally searching a sequence space comparable to that of the ribozyme type has not yet been established for the virus-type.

[0014]

The reason for this is that viruses currently used in the method such as phage displays are parasites of existing cells and are therefore unavoidably subject to restraints

imposed by the host cells, among which can be listed: (1) that only a limited sequence space can be searched owing to restriction by the cells, (2) membrane permeability, (3) bias due to host, and (4) limitation on library owing to host population.

[0015]

The polysome display method (Mattheakis, L. C. & Dower, W. J. (1995) WO95/11922) joins a nucleic acid and a protein via a ribosome by non-covalent bonding. It is therefore suitable when the chain length at the peptide position is short but encounters handling problems when the chain length is long as a protein. Since the huge ribosome remains interposed, the conditions at the time of the selection operation (e.g., adsorption, elution or the like) are subjected to severe restriction. The encoded combinatorial library (Janda, F. H. & Lerner, R. A. (1996) WO96/22391) assigns a chemically synthesized peptide to a nucleic acid tag via beads. Since the yield of chemical synthesis of proteins with around 100 residues is extremely poor using currently available technologies, however, this technique can be used with short chain-length peptides but not with long chain-length proteins.

[0016]

[Problem to be Solved by the Invention]

One conceivable method of overcoming these problems is use of a cell-free translation system. A virus-type strategy molecule that simply binds the genotype and the phenotype in the cell-free systems has a number of advantages including

the following: (1) that a huge mutant population approaching that of the ribozyme-type can be synthesized, (2) creation of various proteins without dependence on a host, (3) no problem regarding membrane permeability, and (4) that the 21st code can be used to introduce a non-native amino acid.

[0017]

An object of the present invention is to provide a molecule comprising a virus-type operation replicon which has the advantages of the aforementioned virus-type strategy molecule, exhibits a higher efficiency than phages, and suffers fewer limitations concerning environmental condition setting, namely, a molecule which should be called "in vitro virus", wherein a nucleic acid and a protein are bound by a chemical bond, that is, a molecule in which a genotype is assigned to a phenotype. More specifically, the present invention has been accomplished in order to provide a molecule exhibiting one-on-one relationship between information and function, which can be utilized for creation of functional proteins and peptides, by performing genotype (nucleic acid) assignment to phenotype (protein) using a cell-free protein synthesis system, and binding the 3'-terminal end of a gene to the C-terminal end of a protein with a covalent bond on ribosome.

[0018]

[Means to Solve the Problem]

The present inventors earnestly conducted investigations to achieve the aforementioned objects, and

as a result, they found that two kinds of molecules that assign a genotype to a phenotype, comprising a nucleic acid and a protein which were chemically bound can be constructed on a ribosome in a cell-free protein synthesis system. The present invention has been accomplished based on these findings.

[0019]

Thus the present invention provides a molecule assigning a genotype to a phenotype, which comprises a nucleic acid portion composed of a polymer of nucleic acid, and a protein portion.

According to preferred embodiments of the present invention, there are provided the aforementioned assigning molecule wherein a 3'-terminal end of the nucleic acid portion and a C-terminal end of the protein portion are bound by a covalent bond, and the aforementioned assigning molecule wherein a 3'-terminal end of the nucleic acid portion covalently bound to a C-terminal end of the protein portion is puromycin.

[0020]

According to another preferred embodiment of the present invention, there is also provided the aforementioned assigning molecule wherein the nucleic acid portion comprises a gene, and the protein portion is a translation product of the gene of the nucleic acid portion; the aforementioned assigning molecule wherein the nucleic acid portion comprises a gene composed of RNA, a spacer and a suppressor tRNA, and

the protein portion is a translation product of the gene of the nucleic acid portion; the aforementioned assigning molecule wherein RNA of the nucleic acid comprises an initiation codon and a termination codon, wherein the suppressor tRNA comprises an anticodon corresponding the termination codon (umber) and wherein the protein portion is a translation product of the RNA gene of the nucleic acid portion; the aforementioned assigning molecule wherein the nucleic acid portion comprises a gene composed of RNA and a spacer portion composed of DNA and RNA, and the protein portion is a translation product of the gene of the nucleic acid portion; and the aforementioned assigning molecule wherein the nucleic acid portion comprises a gene composed of DNA and a spacer portion composed of DNA and RNA, and the protein portion is a translation product of the gene of the nucleic acid portion.

[0021]

As further aspects of the present invention, there are provided a method for constructing an assigning molecule, which comprises (a) bonding a DNA comprising a sequence corresponding to a suppressor tRNA, to a 3'-terminal end of a DNA (gene) through a spacer, (b) transcribing the obtained DNA bonded product into RNA, (c) bonding, to a 3'-terminal end thereof, a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, which can be covalently bonded to an amino acid or a substance having a chemical structure analogous to that of an amino

acid, and (d) performing protein synthesis in a cell-free protein synthesis system using the obtained RNA bonded product as a gene to connect a genotype to a phenotype; and a method for constructing an assigning molecule, which comprises (a) preparing a DNA (gene) which has no termination codon, (b) transcribing the prepared DNA into RNA (gene), (c) bonding a chimeric spacer composed of DNA and RNA to a 3'-terminal end thereof, (d) bonding, to a 3'-terminal end thereof, a nucleic acid or a substance having a chemical structure analogous to that of a nucleoside and a base, which can be covalently bonded to an amino acid or a substance having a chemical structure analogous to that of an amino acid, and (e) performing protein synthesis in an *E. coli* cell-free protein synthesis system using the obtained RNA-DNA chimeric bonded product as a gene to connect a genotype to a phenotype.

[0022]

According to a preferred embodiment of the present invention, there is provided the aforementioned construction method wherein the nucleic acid or the substance having the chemical structure analogous to that of the nucleoside and the base is puromycin.

[0023]

[Embodiment of the Invention]

In this specification, some technical terms are used, and those technical terms have the following meanings when herein used. The term "nucleic acid portion" means a bonded

product of a nucleic acid or a substance having a chemical structure analogous to a nucleic acid and a base, for example, RNA, DNA, PNA (peptide nucleic acid; polymers comprising nucleic acid bases linked via amino acid analogues) and the like, and "protein portion" means a bonded product of an amino acid or a substance having a chemical structure analogous to an amino acid such as naturally-occurring amino acids and non-naturally-occurring amino acids. The term "suppressor tRNA (sup tRNA)" means a tRNA which can suppress mutation by structural change, for example, reading a termination codon on mRNA as a codon corresponding to a certain amino acid.

[0024]

The spacer located at the 3'-terminal end side of the nucleic acid portion may be any spacer provided that it is a polymer substance preferably having a length of 70 to 100Å. Specifically, DNA, RNA, and polymer materials such as polysaccharides, which are naturally-occurring or synthetic, synthetic organic polymer substances such as polyethylene glycols, and the like can be mentioned.

The nucleic acid portion and the protein portion of the assigning molecule of the present invention may be linked through any bond, but they are preferably linked through a chemical bond such as a covalent bond. In particular, preferred are those formed by bonding a nucleic acid or a substance having a chemical structure analogous to a nucleic acid and a base, or a linked product thereof present at the

3'-terminal end of the nucleic acid portion to an amino acid or a substance having a chemical structure analogous to an amino acid present at the C-terminal end of the protein portion via a chemical bond, for example, a covalent bond.
[0025]

For the bonding between the nucleic acid portion and the protein portion, for example, puromycin, 3'-N-aminoacylpuromycin aminonucleoside (PANS-amino acid), which have an amide bond as the chemical bond at the 3'-terminal end of the nucleic acid portion, e.g., PANS-Gly wherein the amino acid portion is glycine, PANS-Val wherein the amino acid portion is valine, PANS-Ala wherein the amino acid portion is alanine, and further PANS-(any of the other amino acids) wherein the amino acid portion is an of the other amino acids, can be utilized. 3'-*N*-Aminoacyl-adenosine aminonucleoside (AANS-amino acid), which comprises as the chemical bond an amide bond formed by dehydration condensation of the amino group of 3'-aminoadenosine and the carboxyl group of an amino acid, for example, AANS-Gly wherein the amino acid portion is glycine, AANS-Val wherein the amino acid portion is valine, AANS-Ala wherein the amino acid portion is alanine, and further AANS-(any of the other amino acids) wherein the amino acid portion is any of the other amino acids, can also be utilized. Those composed of a nucleoside or a nucleoside bound to an amino acid via an ester bond may also be used. Further, any other materials having a binding mode capable of binding a nucleic acid or

a substance having a chemical structure analogous to a nucleic acid and a base and a substance having a chemical structure analogous to an amino acid, can also be utilized.

[0026]

The molecule assigning the genotype to the phenotype of the present invention can be constructed by, for example, (1) a method where the binding of the nucleic acid portion and the protein portion is formed in a site-directed manner, or (2) a method where the bonding of the nucleic acid portion and the protein portion is formed in a non-site-directed manner, which will be explained hereinafter.

First, (1) the method where the bonding of the nucleic acid portion and the protein portion is formed in a site-directed manner will be explained.

[0027]

In this method, a molecule assigning a genotype to a phenotype can be constructed by (a) bonding a DNA comprising a sequence corresponding to sup tRNA, to the 3'-terminal end of a DNA (gene) through a spacer, (b) transcribing the obtained DNA bonded product into RNA, (c) bonding, to the 3'-terminal end thereof, a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, which can be covalently bound to an amino acid or a substance having a chemical structure analogous to that of an amino acid, e.g., puromycin, (d) performing protein synthesis in a cell-free protein synthesis system, e.g., an *E. coli* cell-free protein synthesis system, using the

obtained RNA bonded product as a gene, and thus (e) affording a molecule assigning a genotype to a phenotype comprising a gene RNA (genotype) and a protein (phenotype) which is a translation product of the gene, which are chemically bound through a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, e.g., puromycin.

[0028]

That is, according to this method of the present invention, when a termination codon comes into the A site of ribosome during the protein synthesis, a sup tRNA is correspondingly incorporated, and a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, e.g., puromycin, present at the 3'-terminal end of the sup tRNA is bound to a protein by the action of peptidyl transferase (Figure 2). Therefore, this method is site-directed as for the formation of the bonding between the nucleic acid portion and the protein, which depends on the genetic code.

[0029]

It has been known that puromycin (Figure 3) inhibits the protein synthesis in bacteria (Nathans, D. (1964) Proc. Natl. Acad. Sci. USA, 51, 585-592; Takeda, Y. et al. (1960) J. Biochem. 48, 169-177) and animal cells (Ferguson, J. J. (1962) Biochim. Biophys. Acta 57, 616-617; Nemeth, A. M. & de la Haba, G. L. (1962) J. Biol. Chem. 237, 1190-1193). Puromycin, whose structure resembles the structure of

aminoacyl tRNA, reacts with peptidyl tRNA bound to the P site of ribosome, and it is released from ribosome as peptidyl puromycin, and thus interrupts the protein synthesis (Harris, R. J. (1971) Biochim. Biophys. Acta 240, 244-262).

[0030]

It is not practical to purify native sup tRNA and bond it to mRNA, because of the problems concerning the purification of sup tRNA and the easily hydrolyzable ester bond at the 3'-terminal end of tRNA. Through investigations of tRNA identity, it has been elucidated that unmodified tRNA may be aminoacylated like intact tRNA, and that the aminoacylated unmodified tRNA may be taken into ribosome, and translated (Shimizu, M. et al. (1992) J. Mol. Evol. 35, 436-443). The identity of tRNA was utilized in order to prepare sup tRNA.

[0031]

It has been reported that the aminoacyl synthetases of alanine, histidine, and leucine do not recognize the anticodons thereof (Tamura, K. et al. (1991) J. Mol. Recog. 4, 129-132). Therefore, it can be expected that, by replacing the anticodon of tRNA for alanine with a termination codon of amber, tRNA for alanine (sup tRNA) would be incorporated corresponding to the termination codon.

[0032]

In this respect, it comes into question whether tRNA whose 5'-terminal-end side is not made up by RNAse P or the like, unlike ordinary tRNA, may enter into the A site of

ribosome or not. This is the most important problem to be investigated in determining feasibility of the model of the present invention. It has been known that the 3'-terminal ends of Brome Mosaic Virus (BMV) and Turnip Yellow Mosaic Virus (TYMV) have a tRNA-like structure, and they are aminoacylated by aminoacyl synthetase, and incorporated at an efficiency of 1% in a cell-free translation system (Chen, J. M. & Hall, T. C. (1973) Biochemistry 12, 4570-4574). Supposing that RNA of BMV is incorporated even by 1% by ribosome, it can be expected that RNA having intact tRNA at its 3'-terminal end may be incorporated more efficiently. Even if it is incorporated at an efficiency of 10% or less of that of intact tRNA, there is a reasonable possibility that it can win the competition with the release factor by the concentration effect.

[0033]

Therefore, before the experiment for bonding a protein to the 3'-terminal end of mRNA-sup tRNA (mRNA ligated at its 3'-terminal end with sup tRNA through a spacer), it was examined whether even sup tRNA separated from mRNA entered into the A site of ribosome and was bound to a protein. A sup tRNA whose 3'-terminal end was bonded to puromycin was actually prepared, and added to a cell-free protein synthesis system to examine whether the sup tRNA portion entered into the A site of ribosome corresponding to occurrence of a termination codon and bound to a protein. The 4 repeats region of tau protein (127 residues) was used as mRNA (Goedert,

M. (1989) EMBO J. 8, 392-399). As a result, when the translation was performed in a cell-free protein synthesis system, it could be confirmed that the sup tRNA having puromycin at its 3'-terminal end was incorporated into the A site of ribosome corresponding to a termination codon and bound to a protein (Figure 2).

[0034]

Then, RNA-sup tRNA bonded products having different lengths of the spacer between mRNA and sup tRNA were constructed, and it was attempted to select an optimum length of the spacer which afforded the best efficiency of the incorporation of the sup tRNA portion into the A site of ribosome by the in vitro selection method. As a result, it was found that the RNA-sup tRNA bonded product having a certain spacer length was chemically bound to a protein that was a translation product thereof with a good efficiency.

[0035]

In order to construct the molecule assigning the genotype to the phenotype of the present invention, a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, which is to be bonded to the 3'-terminal end of the nucleic acid portion, and can be covalently bound to an amino acid or a substance having a chemical structure analogous to that of an amino acid, e.g., 2'-deoxycytidylylpuromycin (dCpPur) (Figure 3), must be synthesized first.

[0036]

For synthesizing dCpPur, first, puromycin-5'-monophosphate can be prepared by chemically phosphorylating the 5'-hydroxyl group of puromycin using phosphorus oxychloride and trimethylphosphate. Then, the amino group of the amino acid portion and the 2'-hydroxyl group of the ribose portion in puromycin-5'-monophosphate can be protected by reacting puromycin-5'-monophosphate with trifluoroacetic acid and trifluoroacetic anhydride. The protected product can be reacted with Bz-DMT deoxycytidine in which the amino group of the pyrimidine ring and the 5'-hydroxyl group of the ribose portion in deoxycytidine are protected, in the presence of a condensation agent, dicyclohexylcarbodiimide, and then deprotected with acetic acid and ammonia to afford 2'-deoxycytidylylpuromycin (dCpPur). pdCpPur can be obtained by phosphorylating the 5'-hydroxyl group of dCpPur with polynucleotide kinase.

[0037]

Then, the construction of a bonded product of the nucleic acid portion for binding the nucleic acid portion and the protein portion in a site-directed manner will be described hereinafter.

As the bonded product of the nucleic acid portion used for the site-directed method, for example, a bonded product comprising 5'-(T7 promoter region)-(Shine-Dalgarno (SD) sequence region)-(mRNA region)-(spacer region)-(sup tRNA region)-(puromycin region)-3' connected in this order in sequence can be mentioned.

[0038]

In the construction of this bonded product for the nucleic acid portion, a plasmid comprising the 4 repeats region, which is a microtuble-binding region of human tau protein called htau24 (Goedert, M. (1989) EMBO J. 8, 392-399), inserted downstream of T7 promoter (pAR3040) is constructed first, and it is digested with restriction enzymes *Bgl*II and *Bam*HI to afford a linear DNA. This DNA is used as a template, and amplification is carried out by PCR by using primers for upstream region containing T7 region (forward) and for downstream region containing the SD region and a region around the initiation codon (backward), and Taq DNA polymerase.

[0039]

In the above method, three methionines may be added to the backward primer in order to enhance detection sensitivity for radioactive methionine in the protein portion after the protein synthesis. That is, leucine at position 4, and lysines at positions 5 and 8 of the 4 repeats region are replaced with methionines. Eventually, the translated 4 repeats protein contains four methionines in total. Then, amplification is carried out by PCR using a DNA containing the aforementioned linearized 4 repeats region as a template, a complementary chain of the backward primer mentioned above as a forward primer, and a backward primer which is designed so that the C-terminal end of the 4 repeats region should have an amber codon as the termination codon.

[0040]

The two kinds of DNA fragments amplified by the PCR, namely, the DNA fragment containing the T7 promoter and the SD region and the DNA fragment containing the 4 repeats region are mixed, initially extended without primers, and then amplified by PCR again by using a primer containing the sequence of T7 promoter as the forward primer, and a primer containing a termination codon at the C-terminal of the 4 repeats region as the backward primer.

[0041]

This DNA bonded product (T7 promoter-SD-4 repeats) is ligated to a double-stranded DNA fragment having cohesive ends at the both ends and composed of 17 residues in tandem by using DNA ligase to afford ligation products having different spacer lengths.

After the ligation, the product was fractionated into three fractions (a, b, c) based on the length by polyacrylamide gel electrophoresis (PAGE). The spacer is represented as (17)n wherein n = 15 to 18 for the fraction a, n = 6 to 14 for the fraction b, and n = 0 to 5 for the fraction c. As the sup tRNA, a native alanyl tRNA whose several sites and anticodon are modified into amber (UAG) is prepared by chemical synthesis. This sup tRNA is ligated to the ligation products of the fractions a, b and c having different spacer lengths by using T4 DNA ligase. For the ligation site, an excessive amount of a single-stranded backing DNA is used, and after once melted by temperature elevation, the strands are annealed, and a complementary strand is formed, and

ligated. After the ligation, the ligation product is amplified by PCR using primers for the 5' end and 3'-terminal end of the ligation product. This DNA ligation product is transcribed by using T7 RNA polymerase to form an RNA ligation product.

[0042]

By ligating the pdCpPur chemically synthesized in the above to the 3'-terminal end of this RNA ligation product using T4 RNA ligase, there can be obtained an RNA ligation product, 5'-(T7 promoter region)-(SD region)-(4 repeats region)-(spacer region)-(sup tRNA region)-(puromycin)-3', which can be used as a gene in a cell-free protein synthesis system.

The protein synthesis is performed by adding the above RNA ligation product as a gene to a cell-free protein synthesis system such as cell-free protein synthesis extracts of *E. coli*. In order to obtain an optimum spacer length for obtaining the most efficient bonding of the nucleic acid portion (RNA) and the protein portion, the following experiment is performed.

[0043]

That is, the protein synthesis is performed in a cell-free protein synthesis system (for example, of *E. coli*) by using aforementioned RNA ligation products having three kinds of different spacer lengths, corresponding to the fractions a, b and c, as the gene. In this synthesis, by adding tRNA charged with modified lysine comprising biotin

bound through the ε-amino group of the lysine, biotinyllysine is incorporated in several positions of lysine residues in the translated 4 repeats protein. After the protein synthesis, magnetic beads coated with streptavidin on their surfaces are added to isolate the protein incorporating the biotin.

[0044]

If the nucleic acid portion (RNA) has bonded to the protein portion through puromycin, the nucleic acid portion (RNA) should be bound to the C-terminal end of the protein. When reverse transcription was performed by using a sequence corresponding to the N-terminal region of the 4 repeats as a forward primer and the 3'-terminal end portion of sup tRNA as a backward primer and analyzed by polyacrylamide gel electrophoresis to confirm whether the RNA-protein-bonded product was actually picked up by the magnetic beads, a band of reverse transcribed DNA was observed only for the spacer length of the fraction c. This means that the RNA ligation product having the spacer length of the fraction c is most efficiently bound to the protein portion.

[0045]

Now, it will be explained about (2) the method where the bonding of the nucleic acid portion and the protein portion is formed in a non-site-directed manner.

In this method, a molecule assigning a genotype to a phenotype can be constructed by (a) preparing a DNA containing a gene which has no termination codon, (b) transcribing it

into RNA (gene), (c) bonding a chimeric spacer composed of DNA and RNA to the 3'-terminal end thereof, (d) bonding, to the 3'-terminal end of the bonded product, a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, which can be covalently bound to an amino acid or a substance having a chemical structure analogous to that of an amino acid, e.g., puromycin, (e) performing protein synthesis in a cell-free protein synthesis system, e.g., *E. coli* cell-free protein synthesis system, using the obtained RNA-DNA chimeric bonded product as a gene, and thus (f) affording a molecule assigning a genotype to a phenotype comprising a gene RNA and a protein which is a translation product thereof, which are chemically bound through puromycin.

[0046]

That is, according to this method of the present invention, a nucleic acid or a substance having a chemical structure analogous to that of a nucleic acid and a base, e.g., puromycin, present at the 3'-terminal end of the nucleic acid portion does not enter into the A site of ribosome corresponding to the termination codon of mRNA on ribosome, but randomly enters depending on the spacer length, and puromycin or the like at the 3'-terminal end of the RNA-DNA chimera nucleic acid portion is chemically bound to a protein by the action of peptidyl transferase (Figure 4). Therefore, this method is non-site-directed as for the formation of the bonding between the nucleic acid portion and the protein,

which does not depend on the genetic code.

[0047]

In this method, the molecule assigning the genotype to the phenotype can be constructed by using a non-site-directed ligation product for the nucleic acid portion in the same manner as in the aforementioned site-directed method of (1).

As the ligation product constituting the nucleic acid portion used for the non-site-directed method, for example, a ligation product composed of 5'-(T7 promoter region)-(Shine-Dalgarno (SD) sequence region)-(mRNA region)-(spacer region)-(puromycin region)-3' connected in this order in sequence can be mentioned.

[0048]

In the construction of this ligation product for the nucleic acid portion, the construction from the T7 promoter region to the end of the 4 repeats region may be similar to that explained for the construction of the ligation product for the nucleic acid portion used in (1) the site directed method mentioned above, provided that a primer designed not to have a termination codon by replacing the two termination codons at the C-terminal end of the 4 repeats, ochre (CTG) and amber (TAA), with CAG (glutamine) and AAA (lysine), respectively, is used as a backward primer for the PCR amplification of the ligation product constructed above used as a template.

[0049]

This DNA ligation product is transcribed as a template by using T7 RNA polymerase to afford a corresponding RNA ligation product. This single-stranded RNA ligation product is separately ligated to each of single-stranded chemically-synthesized DNA linkers (chain length; 20, 40, 60, and 80 nucleotides) by using T4 RNA ligase. Then, each ligation product is ligated to a single-stranded DNA-RNA chimeric oligonucleotide comprising 25 residues (DNA; 21 residues, RNA; 4 residues), which is designated as peptide acceptor, by using T4 DNA ligase in the presence of a single-stranded backing DNA.

[0050]

By ligating the pdCpPur chemically synthesized in the above to the 3'-terminal end of the above ligation product using T4 RNA ligase, there can be obtained an RNA-DNA chimeric ligation product, 5'-(T7 promoter region (RNA))-(SD region (RNA))-(4 repeats region (RNA))-(spacer region (DNA))-(peptide acceptor region)-(puromycin)-3', which can be used as a gene for a cell-free protein synthesis system.

[0051]

If protein synthesis is performed by adding the RNA-DNA chimeric ligation product mentioned above as a gene to a cell-free protein synthesis extract (for example, of *E. coli*), there can be obtained a bonded product comprising a nucleic acid portion (RNA-DNA chimeric ligation product) and a protein portion, which are connected by a chemical bond through puromycin.

[0052]

The genetic engineering techniques mentioned in the present specification such as isolation and preparation of nucleic acids, ligation of nucleic acids, synthesis of nucleic acids, PCR, construction of plasmids, and protein synthesis in cell-free system can be performed by the methods described in Sambrook et al. (1989) Molecular Cloning, 2nd Edition, Cold Spring Harbor Laboratory Press, or similar methods unless otherwise indicated.

[0053]

[Effect of the Invention]

Industrial Applicability

According to the present invention, a molecule assigning a genotype (nucleic acid portion) to a phenotype (protein portion), and construction methods therefor are provided. The molecule assigning the genotype to the phenotype of the present invention is an extremely useful substance for evolutionary molecular engineering, i.e., modification of functional biopolymers such as enzymes, antibodies, and ribozymes, and creation of biopolymers having functions which cannot be found in living organisms.

[0054]

[Examples]

The present invention will be more specifically explained with reference to the following examples. However, the following examples should be construed to be an aid for obtaining more specific understanding of the present

invention, and the scope of the present invention is not no way limited by these examples.

[0055]

Example 1:

<1> Preparation of 3'-terminal-end portion of nucleic acid portion

(a) Synthesis of phosphorylated puromycin (pPur)

Materials:

Puromycin (3'-[α-amino-p-methoxyhydro-cinnamamido]-3'-deoxy-*N*,*N*'-dimethyl-adenosine) was purchased from Sigma. Phosphorus oxychloride, and trimethyl phosphate were purchased from Wako Pure Chemicals.

[0056]

Methods:

A solution formed by mixing phosphorus oxychloride (1.5 mmol) and trimethyl phosphate (11.4 mmol) was ice-cooled, and puromycin (0.3 mmol) was added thereto to mix sufficiently, and the mixture was allowed to react at 0°C for 7 hours (Yoshikawa, M. et al. (1969) Bull. Chem. Soc. Jap. 42, 3505-3508). The reaction mixture was then added to an ice-cooled mixture of acetone (40 ml) and ether (20 ml) containing sodium perchlorate ($NaClO_4$, 0.4 g), and stirred sufficiently. Then, water (720 ml) was added to the mixture, and the mixture was stirred at 4°C for 24 hours to hydrolyze the chlorine group. The product precipitated after the hydrolysis was separated by centrifugation, and washed with acetone and ether. The resulting white powder was dried in

vacuo to afford phosphorylated puromycin with a yield of 70 to 90% based on the puromycin.

[0057]

(b) Protection of phosphorylated puromycin by acetylation

Materials:

Trifluoroacetic acid (TFA) was purchased from Nacalai Tesque. Trifluoroacetic anhydride (TFAA) was purchased from Wako Pure Chemicals.

[0058]

Methods:

The dried phosphorylated puromycin (0.2 mmol) and TFA (5 ml) were mixed, and TFAA (2 ml) was added to the mixture at -10°C, followed by stirring. The mixture was allowed to react at room temperature for 1 hour with stirring (Weygand, F. & Gieger, R. (1956) Chem. Ber. 89, 647-652). The reaction was quenched by adding water (50 ml), and the TFA was removed by repeating a procedure comprising addition of water (10 ml) and evaporation to dryness under reduced pressure 5 times. Finally, water (50 ml) was added to the resulting product, followed by lyophilization to afford phosphorylated puromycin in which the amino group of the amino acid portion in puromycin and the 2'-hydroxyl group of the ribose portion were protected with acetyl groups with a yield of 50 to 60% based on the phosphorylated puromycin.

[0059]

(c) Synthesis of dCpPur (2'-deoxycytidyl(3'→5')puromycin)

Materials:

BZ-DMT deoxycytidine (N^4-benzoyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine) was purchased from Sigma, and DCC (dicyclohexyl carbodiimide) was purchased from Watanabe Chemical. The pyridine was purchased from Nacalai Tesque.

[0060]

Methods:

The phosphorylated puromycin protected with acetyl groups (40 μmol) and Bz-DMT deoxycytidine (600 μmol) were dehydrated by repeating a procedure comprising addition of pyridine (2 ml) and evaporation to dryness three times, and finally pyridine (2 ml) was added thereto. DCC (400 μmol) was added to the mixture with stirring, and the mixture was allowed to react at room temperature for 3 days to 2 weeks (Ralph, R. K. et al. (1965) J. Am. Chem. Soc. 87, 5661-5670; and Harris, R. J. et al. (1972) Can. J. Biochem. 50, 918-926). After the reaction, the DMT group was removed by a reaction with 80% acetic acid (5 ml) for two hours. Then, the acetyl group was removed by a reaction with concentrated aqueous ammonia/ethanol (6 ml, volume ratio: 2:1) at 20°C for 2 days. The concentrated aqueous ammonia was removed by evaporation under reduced pressure, and the residue was dissolved in water (40 ml). The resulting solution was applied on a column packed with QAE-Sephadex A-25 (Pharmacia) and adsorbed thereon. Fractions containing the target product were eluted with 0.5 M triethylamine carbonate (TEAB, pH 7.5), lyophilized, and finally separated by HPLC to afford

deprotected dCpPur with a yield of 1 to 5% based on the puromycin.

[0061]

<2> Preparation of nucleic acid portion (in vitro virus genome)

Two kinds of in vitro virus genomes, i.e., (1) one for bonding a nucleic acid portion and a protein portion in a site-directed manner, and (2) one for bonding a nucleic acid portion and a protein portion in a non-site-directed manner, were prepared.

[0062]

Materials:

An *E. coli* cell-free protein synthesis system (*E. coli* S30 Extract System for Linear Templates) was purchased from Promega. T7 RNA polymerase, T4 DNA ligase, T4 DNA kinase, human placenta ribonuclease inhibitor, *Eco*RI, *Bam*HI, and deoxyribonucleotides were purchased from Takara Shuzo. Restriction enzymes *Bst*NI and *Bgl*II were purchased from New England Labs. As for [^{35}S]-methionine, and [γ-^{32}P]-ATP, those from Amersham, and as for Taq DNA polymerase, those from Kurabo and Grainer were used. As for the other biochemical reagents, those from Sigma and Wako Pure Chemicals were used. A plasmid containing the microtuble-binding region of human tau protein (4 repeats) (pAR3040) was prepared by picking up the full length gene of human tau protein from a cDNA library of human brain cloned in λZAPII by PCR, introducing the gene into a plasmid,

amplifying only the 4 repeats region in the plasmid, and introducing the amplified product into a plasmid. As the PCR (polymerase chain reaction) apparatuses, Model PTC-100 (MJ Research) and Model ASTEC PC800 (Astec) were used.

[0063]

(1) Preparation of genome for site-directed bonding

A. Preparation of DNA for mutated 4 repeats portion

1) A plasmid (pAR3040) comprising microtuble region (4 repeats) of human tau protein (Goedert, M. (1989) EMBO J. 8, 392-399) was constructed, and linearized by digestion with restriction enzymes *Bgl*II and *Bam*HI.

[0064]

2) The 4 repeats portion containing the T7 promoter region and the Shine-Dalgarno sequence was amplified by PCR from the genome prepared above. For this amplification, as primers, Left+ (SEQ ID NO: 1) was used for 5′ side, and Right- (SEQ ID NO: 2) for the 3′ side. Right- had such a sequence that the leucine before the ochre termination codon should be mutated into amber termination codon. The PCR conditions were 92°C/30 seconds for denaturation, 65°C/30 seconds for annealing, and 73°C/1 minute for elongation, and this cycle was repeated 30 times.

[0065]

3) Then, the amplified genome was purified, and mutated by utilizing PCR in order to promote the incorporation of methionine and hence enhance detection of radioactive isotope. That is, primers Left- (SEQ ID NO: 3), and Right+ (SEQ ID

NO: 4) containing a region desired to be mutated were synthesized. First, using the DNA of the above 2) as a template, it was amplified by PCR with primers Left+ and Left-, and the amplified DNA was designated as "Left". Amplification by PCR was also performed with primers Right+, and Right-, and the amplified DNA was designated as "Right". After 5% polyacrylamide denatured gel electrophoresis, "Left" and "Right" were excised form the gel, and extracted. The excised Left and Right were first amplified by PCR under the same conditions as mentioned above without primers. Further, 1 μl taken from this reaction mixture was used as a template, and it was amplified by PCR under the same conditions with primers Left+ and Right-. From the above procedure, DNA of the mutated 4 repeats portion in which the number of methionine was increased from one to four was prepared.

[0066]

B. Ligation of alanine suppressor tRNA (Ala-sup tRNA) containing spacers having different lengths to 4 repeats portion

1) The *Bam*H1 site located in the 3'-terminal-end sequence of the above 4 repeats portion obtained in the above A was digested with *Bam*H1. Then, to remove 3'-terminal-end fragment at the *Bam*H1 site, only the 4 repeats portion in 5'-terminal-end sequence was extracted and purified by using QIAquick PCR Purification Kit (QIAGEN).

[0067]

2) The purified product of the above 1), and Spacer-A (SEQ ID NO: 5) which was phosphorylated at 5'-terminal end by T4 kinase were ligated by using T4 DNA ligase while they were backed with Spacer-B (SEQ ID NO: 6).

3) Spacer-C (SEQ ID NO: 7) which was phosphorylated by T4 kinase, and Spacer-B which had a region complementary to Spacer-C were ligated by using T4 DNA ligase. The reaction was performed at 15°C for 2 hours. Then, the product was purified by ethanol precipitation.

[0068]

4) The products of the above sections 2) and 3), Spacer-D (SEQ ID NO: 8), and sup tRNA (SEQ ID NO: 9) which was phosphorylated at 5'-terminal end were dissolved in T4 DNA ligase buffer, denatured at 85°C for 2 minutes, and cooled on ice. After addition of T4 DNA ligase, it was allowed to react at 15°C for 2 hours, and subjected to phenol extraction, and ethanol precipitation.

[0069]

5) The product obtained in the above 4) was used as a template, and amplification was carried out by PCR using the primer Left+, and a primer 3'Pur- (SEQ ID NO: 10) under the conditions of 92°C/30 seconds for denaturation, 65°C/30 seconds for annealing, and 73°C/1 minute for elongation, which cycle was repeated 30 times. The product was subjected to polyacrylamide denatured gel electrophoresis. Three regions A, B and C exhibiting different migration distances were excised, and DNA was extracted from the regions.

[0070]

6) The DNAs extracted from A, B and C in the above 5) and having different lengths were used as templates and amplification were again carried out by PCR under the same conditions, and lengths of the products were determined by electrophoresis, and they were used as template DNA for transcription. As a result, it was found that the numbers of Spacer-C inserted into each product of the fractions were 0-5 for the fraction c, 6-14 for the fraction b, and 15-18 for the fraction a.

[0071]

C. Preparation of RNA genome and ligation of dCpPur

The regions A, B and C obtained in the above B was transcribed into RNA at 37° for 2 hours by using T7 polymerase. Further, the dCpPur obtained in the above <1> Preparation of 3'-terminal-end portion of the nucleic acid portion was phosphorylated in the presence of ATP by using T4 kinase at 15°C for 24 hours, and ligated to the aforementioned transcribed RNA genome by using T4 RNA ligase at 4°C for 50 hours. From this procedure, an RNA genome comprising sup tRNA having puromycin at its 3'-terminal end could be constructed.

[0072]

(2) Preparation of genome for non-site-directed bonding

A. Preparation of DNA and RNA for mutated 4 repeats portion

DNA of the mutated 4 repeats portion was prepared principally the same method as the aforementioned (1) A.

However, the termination codons were eliminated by changing the two termination codons, amber and ochre, to glutamine and lysine, respectively, and a new primer New/Right- (SEQ ID NO: 10) was synthesized in order to make the 3'-terminal-end sequence purine-rich, and used with Left+ for PCR amplification. The amplification by PCR was performed under the conditions of 92°C/30 seconds for denaturation, 65°C/30 seconds for annealing, and 73°C/1 minute for elongation, which cycle was repeated 30 times. This DNA was used as a template to obtain an RNA genome through a reaction at 37°C for 2 hours utilizing T7 polymerase.

[0073]

B. Ligation of Spacers 1 to 4

After a reaction at 36°C for 1 hour with T4 kinase, of Spacer 1 which was a DNA composed of 21 nucleotides (SEQ ID NO: 11), Spacer 2 which was a DNA composed of 40 nucleotides (SEQ ID NO: 12), Spacer 3 which was a DNA composed of 60 nucleotides (SEQ ID NO: 13), or Spacer 4 which was a DNA composed of 80 nucleotides (SEQ ID NO: 14), the RNA obtained in the above A was ligated to the spacers through a reaction at 10°C for 48 hours using T4 RNA ligase.

[0074]

C. Ligation of peptide acceptor (P-Acceptor)

A peptide acceptor (P-Acceptor, SEQ ID NO: 15), which was a chimeric nucleic acid composed of 21-nucleotide DNA and 4-nucleotide RNA, i.e., 25 nucleotides in total, was synthesized in order to enhance the incorporation efficiency

into ribosomes by ligating it at its 3'-terminal end to dCpPur. To phosphorylate the 5'-terminal end of P-Acceptor, it was reacted at 36°C for 1 hour using T4 kinase. Then, the product was backed with Back3' (SEQ ID NO: 16) having a complementary sequence thereto, and ligated to the 3'-terminal end of each of the spacers prepared in the aforementioned B through a reaction at 16°C for 2 hours using T4 RNA ligase. This P-Acceptor was also directly ligated to the 3'-terminal end of A through a reaction at 10°C for 48 hours using T4 RNA ligase, and the product was designated as Non-Spacer genome.

[0075]

D. Ligation of dCpPur

The dCpPur obtained in the above <1> Preparation of 3'-terminal-end portion of nucleic acid portion was phosphorylated by using T4 kinase at 15°C for 24 hours, and ligated to the 3'-terminal end of each of the genomes prepared in the above C by using T4 RNA ligase at 4°C for 50 hours. By this procedure, chimeric RNA genomes comprising puromycin at its 3'-terminal end could be constructed.

[0076]

<3> Optimization of nucleic acid portion

A. Site-directed method

Each of the RNA genomes prepared in the above <2>, (1), which were classified into each of the lengths corresponding to the fractions a, b and c, was translated in 50 μl of *E. coli* cell-free translation system [*E. coli* S30 Extract Systems for Linear Templates (Promega)] containing

biotinylated lysyl tRNA (Promega), and after addition of 5 mg of streptavidin coated magnetic beads (Dynabeads, Dynal) to each reaction tube, it was incubated at room temperature for 1 hour. Then, the Dynabeads were collected by a magnet, and the supernatant was aspirated. The remained Dynabeads were washed 2 times with B&W buffer (1000 μl). The beads were further washed twice with RT-PCR buffer (500 μl), and resuspended in RT-PCR buffer (500 μl). The suspension (50 μl) was transferred into a 500-μl Eppendorf tube, and after the Dynabeads were immobilized with a magnet, the supernatant was aspirated. To the remained Dynabeads, RT-PCR buffer, reverse transcriptase, and Taq polymerase [Access RT-PCR System (Promega)] were added. Reverse transcription was performed at 48°C for 1 hours, and PCR was performed under the conditions of 94°C/30 seconds, 65°C/40 seconds, and 68°C/1 minute and 40 seconds, which cycle was repeated 40 times, using primers of Right+ (SEQ ID NO: 4) and 3'Pur- (SEQ ID NO: 10). The results of the analysis of the fractions a, b and c by electrophoresis were shown is Figure 5.
[0077]

A band was detected from the group of the fraction c (Lane 3 in Figure 5). This band was excised from the gel by electrophoresis, ligated to "Left" having the T7 promoter and the Shine-Dalgarno region by PCR, and amplified by PCR using the primers Left+ (SEQ ID NO: 3) and 3'Pur- (SEQ ID NO: 10). This genome was designated as "Stranger".
[0078]

Then, to examine whether the protein actually translated from Stranger was bonded to the mRNA portion (RNA genome portion), after transcription, it was ligated with pdCpPur at its 3'-terminal end using T4 RNA ligase, dephosphorylated at 5'-terminal end of RNA using HK phosphatase (Epicentre) at 30°C for 1 hour, and labeled with [γ-^{32}P]-ATP using T4 kinase. The product was added to an *E. coli* cell-free translation system as mRNA, and allowed to react at 37°C for 1 hour and 40 minutes. The results of 18% SDS-PAGE of the product are shown in Figure 6. From the results, it can be seen that the nucleic acid portion and the protein portion were bonded to form in vitro viruses, at a rate of about 80% or more.

[0079]

B. Non-site-directed method

Because a short spacer was already used in the site-directed method, dCpPur, which had been phosphorylated at 5'-terminal end in the presence of [γ-^{32}P]-ATP using T4 kinase, was ligated to the 3'-terminal end of "Non-spacer" RNA genome without a spacer through a reaction using T4 RNA ligase at 4°C for 50 hours. The product was added to an *E. coli* cell-free translation system together with mRNA encoding the ordinary 4 repeats, and allowed to react at 37°C for 1 hour and 30 minutes. This reaction mixture (10 μl) digested with ribonuclease T2, and an equal amount of the reaction mixture were electrophoresed by 18% SDS-PAGE, and analyzed by an image analyzer BAS2000 (Fujifilm) (Figure 7).

[0080]

As a result, as for the sample digested with ribonuclease T2, a band appeared at the same migration distance as that of the control, the 4 repeats labeled with [^{35}S]-methionine, because the sample contained the released protein portion. On the other hand, as for the sample not treated, a band appeared above the 4 repeats protein, i.e., it was found that the band reflected a clearly larger molecular weight. This band did not correspond to the labeled mRNA itself (about 400 nucleotides), because it migrated a longer distance than the tRNA. Therefore, it was identified as a substance composed of bonded RNA and protein. That is, these results demonstrated that the nucleic acid portion was bonded to the protein portion in a non-site-directed manner.

[0081]

[Sequence Listing]

SEQ ID NO:1
LENGTH: 33
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: T7 promoter upstream
SEQUENCE DESCRIPTION:
GAGCATAGAT CTCGATCCCG CGAAATTAAT ACG 34

[0082]

SEQ ID NO:2
LENGTH: 33
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: includes a termination codon

SEQUENCE DESCRIPTION:
GCAGCCGGAT CCTTACTACT TGTGGGTTTC AAT 33

[0083]

SEQ ID NO:3
LENGTH: 33
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: includes an initiation codon; complementary to SEQ ID NO: 4
SEQUENCE DESCRIPTION:
GGACATGACA TTCATCATGT CTGGCATATG TAT 33

[0084]

SEQ ID NO:4
LENGTH: 33
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: includes an initiation codon; complementary to SEQ ID NO: 3
SEQUENCE DESCRIPTION:
ATACATATGC CAGACATGAT GAATGTCATG TCC 33

[0085]

SEQ ID NO:5
LENGTH: 16
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a portion complementary to SEQ ID NO: 6
SEQUENCE DESCRIPTION:
GATCTATTTC TTATTC 16

[0086]

SEQ ID NO:6
LENGTH: 17
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has an initiation codon; complementary to SEQ ID NO: 5
SEQUENCE DESCRIPTION:
GAAGAGAATA AGAAATA 17

[0087]

SEQ ID NO:7
LENGTH: 17
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a portion complementary to SEQ ID NO: 6
SEQUENCE DESCRIPTION:
TCTTCTATTT CTTATTC 17

[0088]

SEQ ID NO:8
LENGTH: 30
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a portion complementary to SEQ ID NO: 9
SEQUENCE DESCRIPTION:
GGGTAAACGA ATGAACAAGA ATAAGAAATA 30

[0089]

SEQ ID NO:9
LENGTH: 108
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a sequence of an alanyl tRNA
SEQUENCE DESCRIPTION:
TTGTTCATTC GTTTACCCGG GGCTATAGCT CAGCTGGGAG AGCGCCTGCT TCTAACGCAG 60
GAGGTCTGCG GTTCGATCCC GCGTAGCTCC ACCAGGAGGC GACTAGCT 108

[0090]

SEQ ID NO:10
LENGTH: 23
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a 3'-side sequence of an alanyl tRNA
SEQUENCE DESCRIPTION:
GTGGAGCTAC GCGGGATCGA ACC 23

[0091]

SEQ ID NO:11
LENGTH: 25
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has no initiation codon
SEQUENCE DESCRIPTION:
GCAGCCGGAT CCTTTCTGCT TGTGG 25

[0092]

SEQ ID NO:12
LENGTH: 21
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a sequence partly complementary to SEQ ID NO: 17
SEQUENCE DESCRIPTION:
CTTTAATGAC CTCCCCTCTC C 21

[0093]

SEQ ID NO:13
LENGTH: 40
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a sequence partly complementary to SEQ ID NO: 17
SEQUENCE DESCRIPTION:
CTTTAATAAT TTTTTTTTTT TTTAATGACC TCCCCTCTCC 40

[0094]

SEQ ID NO:14
LENGTH: 60
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a sequence partly complementary to SEQ ID NO: 17
SEQUENCE DESCRIPTION:
CTTTAATAAT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTAATGACC TCCCCTCTCC 60

[0095]

SEQ ID NO:15
LENGTH: 80
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a sequence partly complementary to SEQ ID NO: 17
SEQUENCE DESCRIPTION:
CTTTAATAAT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT TTTTTTTTTT 60
TTTAATGACC TCCCCTCTCC 80

[0096]

SEQ ID NO:16
LENGTH: 25
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: small capiptals are RNA
SEQUENCE DESCRIPTION:
CTTACTGTCT TTTTTTTTTT Tgagc 25

[0097]

SEQ ID NO:17
LENGTH: 33
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid synthetic DNA
OTHER INFORMATION: has a sequence partly complementary to SEQ ID NO: 16
SEQUENCE DESCRIPTION:
AAAAAAGACA GTAAGGGAGA GGGGAGGTCA TTA 33

[Brief Description of the Drawings]

[Figure 1]

It shows strategies for genotype (nucleic acid portion) assignment to phenotype (protein portion).

[Figure 2]

It shows a method for construction of the molecule assigning the genotype to the phenotype of the present invention wherein a nucleic acid portion and a protein portion are bonded in a site-directed manner.

[Figure 3]

It shows chemically-modified portions of the 3'-terminal ends of nucleic acid portions, which are a point of the construction of the molecule assigning the genotype to the phenotype (in vitro virus).

[Figure 4]

It shows a method for construction of the molecule assigning the genotype to the phenotype of the present invention wherein a nucleic acid portion and a protein portion are bonded in a non-site-directed manner.

[Figure 5]

It is a photograph of electrophoresis image that shows spacer optimization in the site-directed method. It shows the results of 4% polyacrylamide gel electrophoresis (in the presence of 8 M urea) of a DNA obtained through a process comprising translation of each RNA genome having a spacer in a length corresponding to each of the prepared fractions a, b, and c in the presence of a biotinylated lysyl tRNA in an *E. coli* cell-free translation system, specific absorption on streptavidin-coated magnetic beads, reverse

transcription, and amplification by PCR (staining was silver staining). Lane 1 is for the spacer length of fraction a (255-306 residues), Lane 2 is for the spacer length of fraction 2 (102-238 residues), and Lane 3 is for the spacer length of fraction c (0-85 residues).

[Figure 6]

It is a photograph of electrophoresis image showing bonding of a nucleic acid portion and a protein portion in a site-directed method. The results were obtained by 18% polyacrylamide gel electrophoresis (in the presence of 8 M urea and SDS): Lane 1 for a translation product of mRNA encoding the 4 repeats region of a tau protein, which was obtained in an *E. coli* cell-free translation system while labeled with [^{35}S]-methionine, and Lane 2 for a translation product of the mRNA whose 3'-terminal end was bonded to sup tRNA having puromycin, and whose 5'-terminal end was labeled with [^{32}P], which was obtained in an *E. coli* cell-free translation system.

[Figure 7]

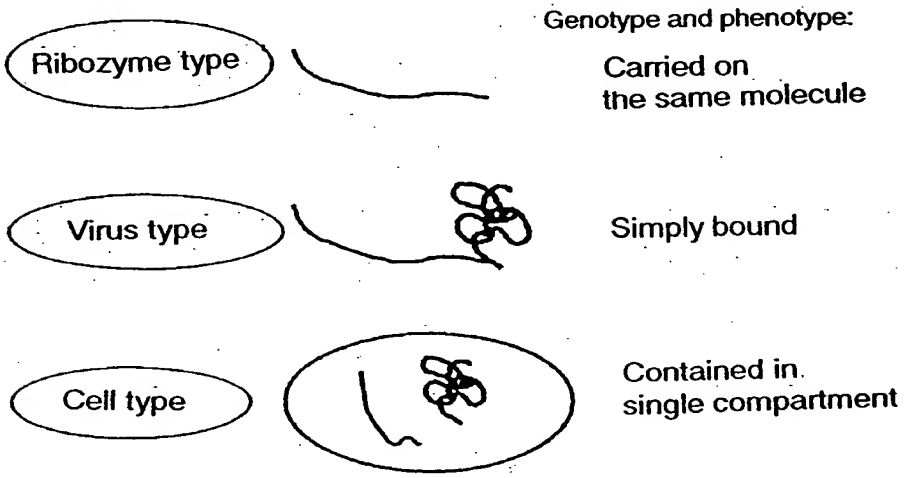
It is a photograph of electrophoresis image showing bonding of nucleic acid portion and protein portion in the non-site-directed method. The results were obtained by 18% polyacrylamide gel electrophoresis (in the presence of SDS): Lane 1 for a translation product of mRNA encoding the 4 repeats region of a tau protein, which was obtained in an *E. coli* cell-free translation system, while labeled with [^{35}S]-methionine, and Lane 2 for a translation product of

the mRNA whose 3'-terminal end was bonded through a spacer to puromycin labeled with [^{32}P] at the 5′ end, which was obtained in an *E. coli* cell-free translation system, and Lane 3 for the translation product of Lane 2 digested with ribonuclease T2.

[Name of Document] Drawings

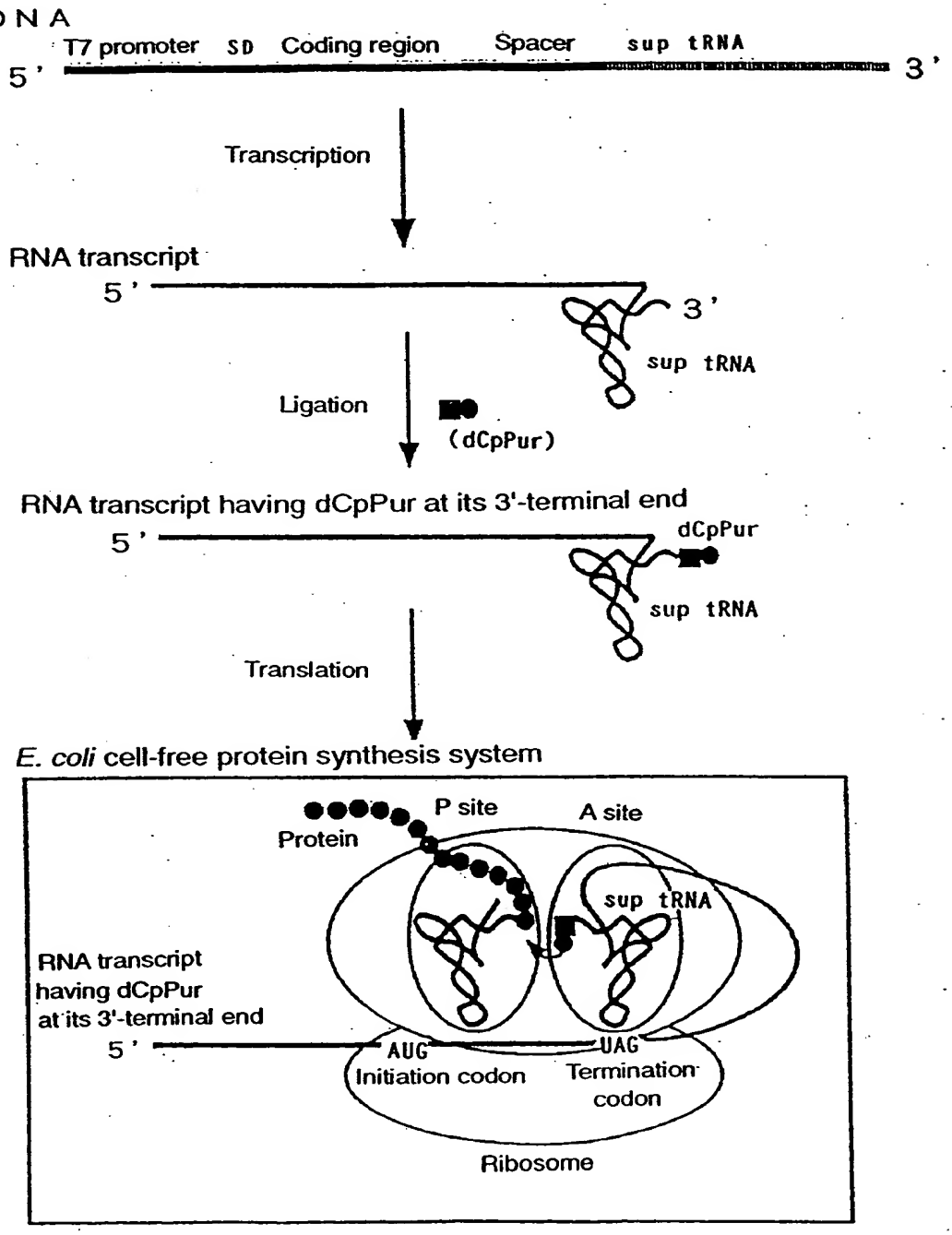
[Fig. 1]

Strategy for assigning genotype to phenotype



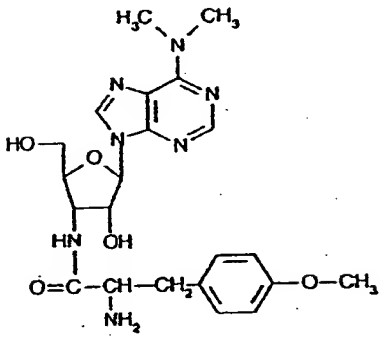
Strategies for assignment are logically classified into three patterns.

[Fig. 2]

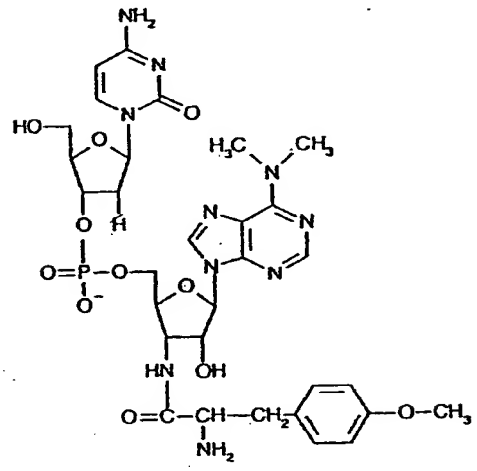


[Fig. 3]

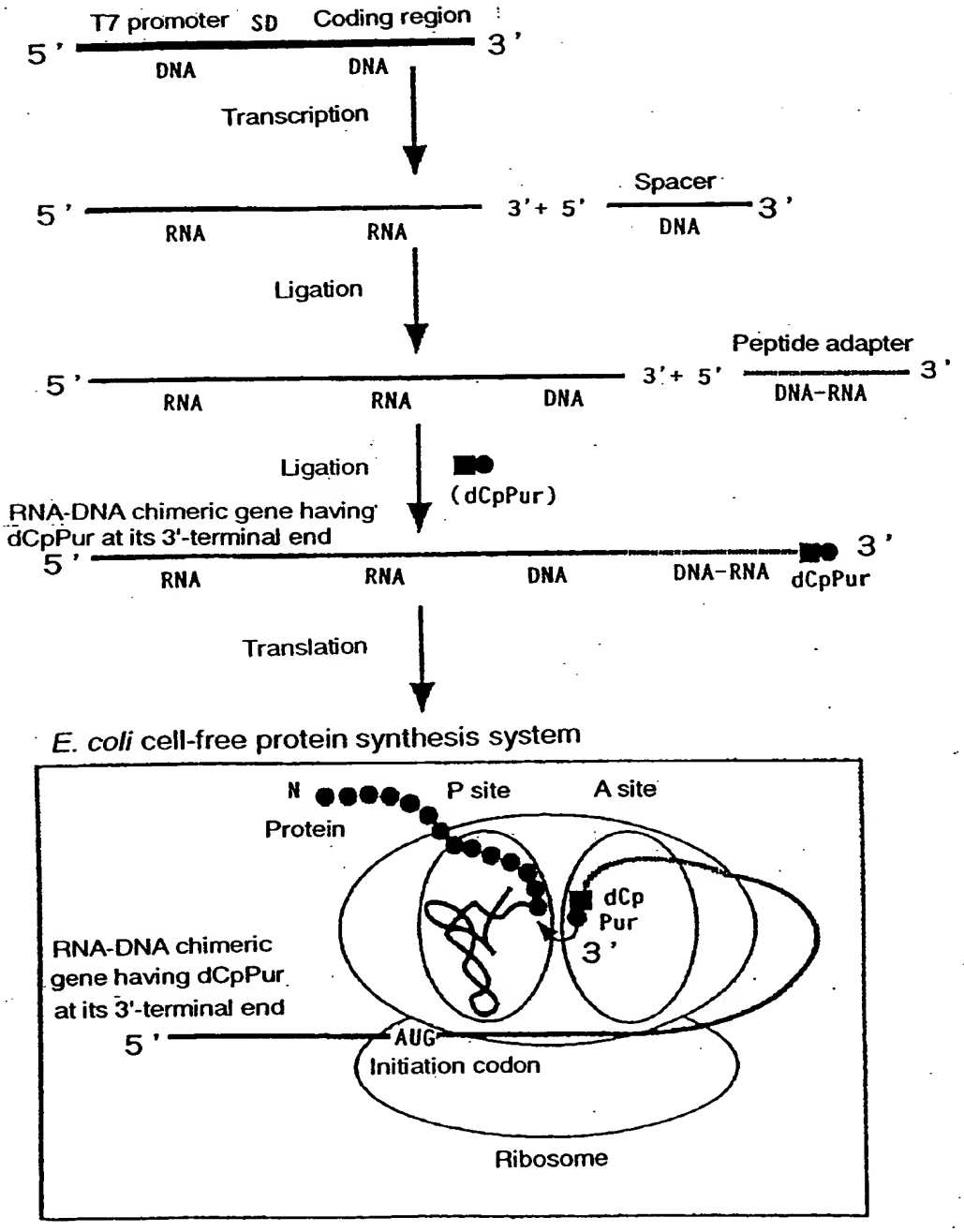
Puromycin (Pur)



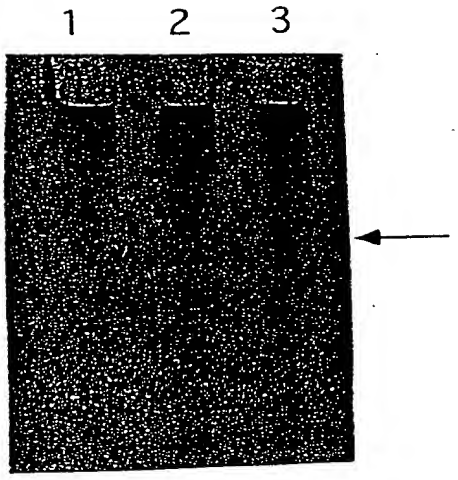
2'-Deoxycytidylpuromycin (dCpPur)



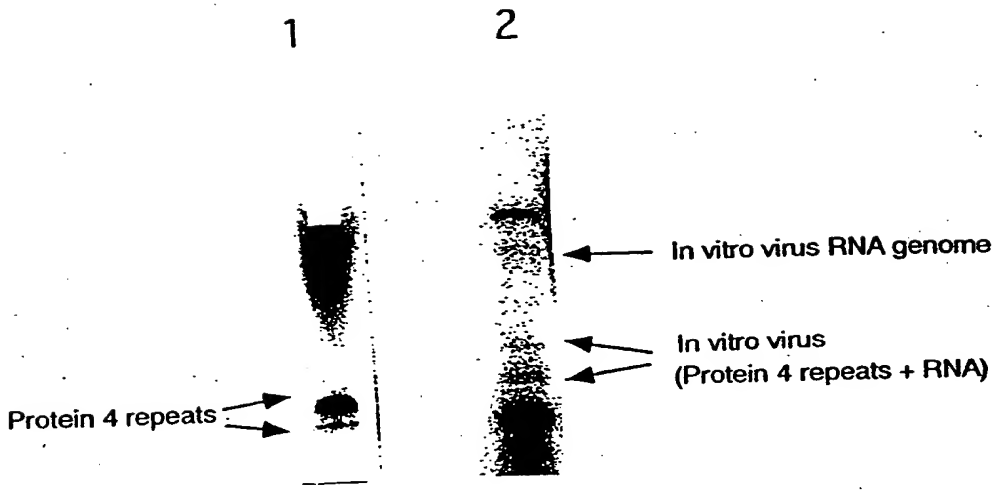
[Fig. 4]



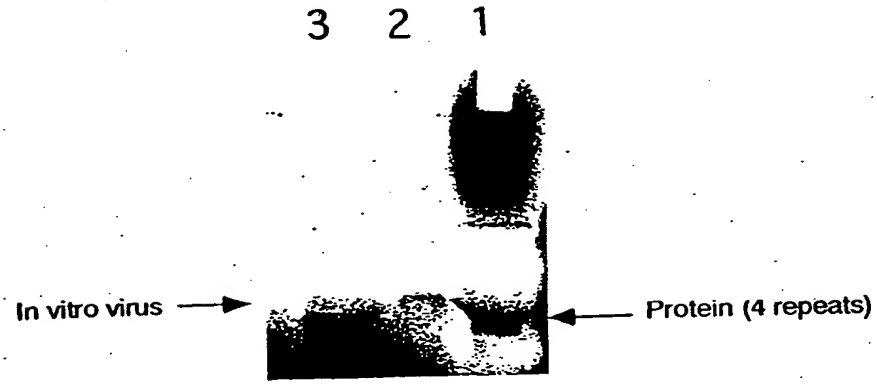
[Fig. 5]



[Fig. 6]



[Fig. 7]



ABSTRACT OF THE DISCLOSURE

A molecule comprising a nucleic acid portion and a protein portion directly bound to said nucleic acid portion with a covalent bond, wherein said nucleic acid portion comprises a polymer of nucleoside, and said protein portion is encoded by said nucleic acid portion, and a method for constructing the molecule, which comprises (a) preparing a DNA containing a gene which has no termination codon, (b) transcribing the prepared DNA into RNA, (c) bonding a chimeric spacer composed of DNA and RNA to a 3'-terminal end of the obtained RNA, (d) bonding to a 3'-terminal end of the obtained bonded product, a nucleoside or a substance having a chemical structure analogous to that of a nucleoside, which can be covalently bound to an amino acid or substance having a chemical structure analogous to that of an amino acid, and (e) performing protein synthesis in a cell-free protein synthesis system using the obtained bonded product as mRNA to bond a nucleic acid portion containing the gene to a translation product of the gene. The molecule of the present invention is an extremely useful substance that can be used for evolutionary molecular engineering, i.e., modification of functional biopolymers such as enzymes, antibodies, and ribozymes, and creation of biopolymers having functions which cannot be found in living organisms.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of	:	
Hiroshi YANAGAWA et al.	:	Docket No. 01416/OP623-PC-US
Serial No. 09/284,627	:	Group Art Unit 1636
Filed June 2, 1999	:	Examiner A. Ousley
MOLECULE ASSIGNING GENOTYPE TO PHENOTYPE AND USE THEREOF	:	

AMENDMENT

Assistant Commissioner for Patents,
Washington, D.C. 20231

THE COMMISSIONER IS AUTHORIZED TO CHARGE ANY DEFICIENCY IN THE FEE FOR THIS PAPER TO DEPOSIT ACCOUNT NO. 23-0975.

Sir:

Responsive to the Official Action dated May 24, 2000, the time for filing thereto being extended for three months in accordance with the Petition for Extension submitted concurrently herewith, please amend the above-identified application as follows.

In the Specification:

Page 1, between lines 2 and 4, insert the heading --Background of the Invention--;

line 4, change the heading to read --1. Field of the Invention--.

Page 4, line 24, change the heading to read --2. Description of the Related Art--.

Page 9, line 9, change the heading to read --Summary of the Invention--.

Page 10, line 22, after "Thus" insert a comma --,--.

Page 21, line 24, change the heading to read --Description of the Preferred

Embodiments--.

Page 67, line 11, delete in its entirety.

In the Abstract:

Please replace the Abstract of record with the substitute Abstract enclosed herewith.

In the Claims:

Kindly cancel claim 4 without prejudice.

Please amend the claims as follows:

1. (Amended) A molecule [assigning a genotype to a phenotype, which comprises] comprising a nucleic acid portion [having a nucleotide sequence reflecting the genotype, and a protein portion comprising a protein involved in exhibition of the phenotype, the nucleic acid portion and the protein portion being directly bound by a chemical bond] and a protein portion directly bound to said nucleic acid portion with a covalent bond, wherein said nucleic acid portion comprises a polymer of nucleoside, and said protein portion is encoded by said nucleic acid portion.

Claim 2, line 1, delete "assigning".

Claim 3, line 1, delete "assigning".

Claim 5, line 1, delete "assigning"; change "claim 4" to --claim 1--.

Claim 6, line 1, delete "assigning".

Claim 7, line 1, delete "assigning"; change "claim 4" to --claim 1--.

Claim 8, line 1, delete "assigning"; change "claim 4" to --claim 1--.

Claim 9, line 1, delete "assigning"; change "claim 4" to --claim 1--.

Claim 10, line 1, delete "assigning"; change "claim 4" to --claim 1--.

Claim 11, line 1, delete "assigning"; change "claim 4" to --claim 1--.

Claim 12, line 1, delete "assigning".

Claim 13, line 1, delete "assigning".

Claim 15, line 1, delete "assigning".

Claim 16, line 1, delete "assigning".

Claim 17, line 1, delete "assigning".

Claim 18, line 2, delete "assigning";

line 5, delete "assigning";

line 8, delete "assigning".

Claim 20, line 3, delete "assigning";

line 6, delete "assigning";

Claim 21, line 1, change "the" to --a--; before "evolution" insert --protein--;

line 2, delete "as defined in claim 18 or 19";

line 3, delete "assigning";

line 19, delete "assigning";

line 37, delete "assigning";

line 39, change "an assigning" to --a--.

Please add the following new claims:

12 ~~22~~. The molecule according to claim 2, wherein the covalent bond is formed by a cell-free protein synthesis system.

23 ~~23~~. The method for protein evolution simulation according to claim 20 ~~18~~, wherein the selection step is conducted using a target substance which is bound to a solid-state surface.

16 ~~24~~. The molecule according to claim 1 or 2, wherein a 3' -terminal end of the nucleic acid portion covalently bonded to a C-terminal end of the protein portion is a substance having the ability to bind to the C-terminal of a synthesized protein when protein synthesis is carried out in a cell-free protein synthesis system.

13

11 ~~25~~. The molecule according to claim 1 or 2, wherein a 3' terminal end of the nucleic acid portion covalently bonded to a C-terminal end of the protein portion is 3'-*N*-aminoacylpuromycin aminonucleoside or 3'-*N*-aminoacyladenosine aminonucleoside.--.

REMARKS

Favorable reconsideration is respectfully requested in view of the foregoing amendments and the following remarks.

With regard to the Applicants' claim for foreign priority, Applicants wish to note that this application is a national stage application under 35 USC § 371 and thus, Applicants' claim for foreign priority should be under 35 USC § 119(a)-(d) and not 35 USC § 119(e) as noted in the

11

Official Action Summary. Applicants respectfully request correction in this regard. Applicants also request acknowledgment of receipt of the certified priority document from the International Bureau as noted in the Notification of Acceptance of Application under 35 USC § 371.

With respect to the objections to the drawings, Applicants respectfully request that the objections be held in abeyance until after an indication of allowable subject matter.

With respect to the Abstract of the Disclosure, Applicants have submitted herewith a new Abstract as required by the Examiner. The new Abstract contains a single paragraph of 220 words and 15 lines of text. Applicants believe that the new Abstract is in compliance with MPEP § 608.01(b) and thus, this objection should be withdrawn.

The specification has been reviewed and editorial changes have been effected. All of the changes are minor in nature and therefore do not require extensive discussion. Specifically, the specification headings have been amended in conformance with U.S. practice.

Claim 4 has been cancelled without prejudice, claims 1-3, 5-13, 15-18, 20 and 21 have been amended, and new claims 22-25 have been added. Support for the claim amendments and new claims is readily apparent from the teachings of the specification and the original claims. Specifically, the term "nucleic acid portion" is defined on page 21, line 27 to page 22, line 4, of the specification. Further, the phrase "protein portion is encoded by the nucleic acid portion" is supported by the description on page 24, line 15 to page 25, line 16 and page 33, line 22 to page 34, line 26 of the specification. From the methods described therein, it is clear that the protein portion is a resultant product of translation of the nucleic acid portion by a protein synthesis

system. Lastly, the phrase "protein portion directly bound to said nucleic acid portion with a covalent bond" is described on page 23, line 5 to page 24, line 14 of the specification.

With regard to the new claims, support for claim 22 can be found on page 33, line 22 to page 34, line 13 of the specification. For claims 23, support is based on the description on page 41, lines 9-21. One skilled in the art can clearly conclude that the "solid-state surface" is inherently disclosed based on the description of "microplate, beads or the like". For claim 24, support is based on the description on page 24, line 15 to page 25, line 16 and page 33, line 22 to page 34, line 26 of the Specification. Applicants note that from the construction method, it is clear that the 3'-terminal end of the nucleic acid portion is a substance having the specified ability. Lastly, support for claim 25 can be found on page 23, line 16 to page 24, line 7 of the specification.

With regard to the objection of claims 4-21 under 37 CFR 1.75(c), this objection has been overcome by the amendments to the claims. Specifically, claim 4 has been cancelled, claims 5 and 7-11 have been amended to depend on claim 1, and claim 21 has been amended to an independent claim. Applicants believe that the claim amendments put the claims in proper multiple dependent form. Thus, the objection of claims 4-21 under 37 CFR 1.75(c) can no longer be sustained and should be withdrawn.

With regard to the rejection of claims 1-3 under 35 USC § 112, second paragraph, this rejection has been overcome by the wording of the amended claims. Specifically, the term "assigning" and the phrases "assigning a genotype to a phenotype" and "a nucleotide sequence

reflecting the genotype, and a protein portion comprising a protein involved in exhibition of the phenotype" have been deleted from the claims.

The Examiner also states that the disclosure of the specification is inconsistent with the accepted and well understood meanings of "phenotype" and "genotype". However, Applicants believe that the Examiner is mistaken in this regard. Applicants believe that the Examiner has misunderstood the teachings of the claims. The molecule in which the specified nucleic acid portion is bound to the specified protein portion, is used for the assignment of the genotype to the phenotype. In other words, the term "phenotype" is not used to mean the protein portion or the protein, and the term "genotype" is not used to mean the nucleic acid portion or the gene. Instead, the terms have been used consistent with the meanings set forth by the Examiner. The protein portion is representative of "a visible or other measureable property being observe in the cell". The nucleic acid portion is "the genetic factor responsible for creating" the protein portion since the nucleic acid portion encodes the protein portion.

With regard to the rejection of claims 1 and 2 under 35 USC § 102(b) as being anticipated by Tung et al., this rejection is deemed to be untenable and is thus respectfully traversed.

To constitute anticipation of the claimed invention, a single prior art reference must disclose each and every material element of the claim. Here, in this case, the molecule of the present invention is a molecule in which a protein is directly bound via a covalent bond to a nucleic acid encoding said protein. Conversely, the oligonucleotide covalently linked to a peptide disclosed in Tung et al. does not encode said peptide. In other words, although Tung et al. disclose a molecule in which a peptide and an oligonucleotide are covalently linked, the peptide

is not encoded by the oligonucleotide. In Tung et al., the peptide is a Tat peptide representing the TAR-binding domain of Tat protein, and the oligonucleotide is an antisense oligonucleotide complementary to the loop region of TAR RNA. The peptide in the conjugate interacts with a folded domain in HIV-1 TAR RNA which encodes the peptide and the oligonucleotide in the conjugate hybridizes with the TAR RNA (see the abstract, lines 2-3, and page 292 of the reference). The TAR RNA is not covalently linked to the peptide encoded by the TAR RNA.

Thus, it is clear from the above that the molecule taught in Tung et al. is distinguishable from the molecule of the present invention. As a result, Applicants respectfully submit that this rejection can no longer be sustained and should be withdrawn.

With regard to the rejection of claims 1 and 2 under 35 USC § 102(b) as being anticipated by Brenner et al., this rejection is also deemed to be untenable and is thus respectfully traversed

Brenner et al. disclose a process of combinatorial chemistry based on a split method in which elongation of oligonucleotide sequences and peptides on solid-state surfaces, and split of the elongated product are repeated (see page 5382, right column, lines 1-23 from the bottom, Steps 1 to 5) . In the molecule obtained by this process, a peptide is bound to a nucleotide (genetic tag) encoding the peptide through the solid-state surface. However, for encoding the peptide, sextuplets are used for convenience's sake, and triplet codons according to the genetic code are not used (see page 5382, right column, see "A Formal Example") which is clearly distinguishable from the molecule of the present invention.

In addition, in the molecule disclosed by Brenner et al., the peptide is bound to the genetic tag through the solid-state surface. Conversely, the nucleic acid portion and the protein

portion of the molecule of the present invention are directly bound to each other via a covalent bond.

Thus, in view of the above, this rejection under 35 U.S.C. § 102(b) can not be sustained and should be withdrawn.

In view of the foregoing amendments and remarks, it is respectfully submitted that the Application is now in condition for allowance. Such action is thus respectfully solicited.

If, however, the Examiner has any suggestions for expediting allowance of the application or believes that direct communication with Applicants' attorney will advance the prosecution of this case, the Examiner is invited to contact the undersigned at the telephone number below.

Respectfully submitted,

Hiroshi YANAGAWA et al.

By: _____

Lee Cheng
Registration No. 40,949
Attorney for Applicants

LC/gtn
Washington, D.C.
Telephone (202) 721-8200
Facsimile (202) 721-8250
November 24, 2000

世界知的所有権機関
国際事務局

PCT 特許協力条約に基づいて公開された国際出願

(51) 国際特許分類6 C12N 15/11, C12P 21/00, C12Q 1/68	A1	(11) 国際公開番号	WO98/16636
		(43) 国際公開日	1998年4月23日(23.04.98)

(21) 国際出願番号 PCT/JP97/03766

(22) 国際出願日 1997年10月17日(17.10.97)

(30) 優先権データ
特願平8/274855 1996年10月17日(17.10.96) JP

(71) 出願人（米国を除くすべての指定国について）
三菱化学株式会社
(MITSUBISHI CHEMICAL CORPORATION)[JP/JP]
〒100 東京都千代田区丸の内二丁目5番2号 Tokyo, (JP)
(72) 発明者；および
(75) 発明者／出願人（米国についてのみ）
柳川弘志(YANAGAWA, Hiroshi)[JP/JP]
根本直人(NEMOTO, Naoto)[JP/JP]
〒194 東京都町田市南大谷11号
株式会社 三菱化学生命科学研究所内 Tokyo, (JP)
宮本悦子(MIYAMOTO, Etsuko)[JP/JP]
〒230 神奈川県横浜市鶴見区寺谷1-18-18 Kanagawa, (JP)
伏見 譲(FUSIMI, Yuzuru)[JP/JP]
〒338 埼玉県浦和市神田671-6 Saitama, (JP)

(74) 代理人
弁理士 遠山 勉, 外(TOYAMA, Tsutomu et al.)
〒103 東京都中央区東日本橋3丁目4番10号
ヨコヤマビル6階 Tokyo, (JP)

(81) 指定国 CA, JP, US, 欧州特許 (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

添付公開書類
国際調査報告書

(54)Title: **MOLECULE THAT HOMOLOGIZES GENOTYPE AND PHENOTYPE AND UTILIZATION THEREOF**

(54)発明の名称 遺伝子型と表現型の対応付け分子及びその利用

(57) Abstract

A molecule that homologizes a genotype and a phenotype and is prepared by covalently joining the 3' end of a nucleic acid portion having a base sequence reflecting a genotype to the C-terminus of a protein portion containing a protein participating in phenotypie expression; and a method of constructing a molecule that homologizes a genotype and a phenotype, which comprises the steps of (a) preparing a DNA containing a gene free from termination codon, (b) transcribing the DNA into a RNA, (c) joining a spacer comprising a DNA-RNA chimera to the 3' end of the RNA, (d) further joining to the 3' end of the product of ligation a nucleoside or a substance having a chemical structure analogous to that of the nucleoside, the nucleoside and the substance analogous thereto being capable of forming a covalent bond with an amino acid or a substance having a chemical structure analogous to that of the amino acid, and (e) synthesizing a protein in a cell-free protein synthesis system by using the product of ligation as a mRNA to thereby join the gene-containing nucleic acid portion to the product of gene translation. This molecule is extremely useful in evolutionary molecular engineering for the modification of functional biopolymers, such as enzymes, antibodies and

a...PROMOTER
b...CODING REGION
c...TRANSCRIPTION
d...SPACER
e...LIGATION
f...PEPTIDE ADAPTER
g...RNA-DNA CHIMERA GENE BEARING dCpPur AT THE 3'END
h...TRANSLATION
i...E. COLI CELL-FREE PROTEIN SYNTHESIS SYSTEM
j...PROTEIN
k...P SITE
l...A SITE
m...INITIATION CODON
n... RIBOSOME